WCMC Handbooks on Biodiversity Information Management

A DARWIN INITIATIVE PROJECT

WORLD CONSERVATION MONITORING CENTRE

COMMONWEALTH SECRETARIAT

Data Management Fundamentals

Volume 7

WCMC Handbooks on Biodiversity Information Management

Volume 7
Data Management Fundamentals

World Conservation Monitoring Centre

WORLD CONSERVATION
MONITORING CENTRE

Series Editor J.H. Reynolds

Commonwealth Secretariat 1998

The **World Conservation Monitoring Centre**, based in Cambridge, UK, is a joint venture between three partners in the *World Conservation Strategy* and its successor *Caring for the Earth*: IUCN – The World Conservation Union, UNEP – United Nations Environment Programme, and WWF – World Wide Fund for Nature. The Centre provides information services on the conservation and sustainable use of species and ecosystems and supports others in the development of their own information systems.

The United Kingdom's **Darwin Initiative for the Survival of Species**, launched at the 1992 Earth Summit in Rio de Janeiro, aims to support the Convention on Biological Diversity by drawing on Britain's scientific, educational and commercial strengths to assist in the conservation and sustainable use of the world's biodiversity and natural habitats. Key tenets of the Darwin Initiative include collaboration and cooperation with local people, capacity building, distinctiveness and complementarity of project initiatives, poverty alleviation, and long-term sustainability. Through training, awareness raising, and research on undervalued areas of biodiversity, Darwin support is particularly aimed at strengthening links between Britain and those countries rich in biodiversity but poor in financial resources.

Under the auspices of its **Environmental Training for Sustainable Development** initiative, the Management and Training Services Division of the **Commonwealth Secretariat** supports short- and long-term training, internships and institution development for environmental policy makers, environmental 'operatives', and environmental information professionals in the Commonwealth, in various areas of the environment including biodiversity and gender. Funding support for training, institution development and publications under the aegis of the Management and Training Services Division is provided by the Fund for Technical Co-operation (CFTC).

WORLD CONSERVATION MONITORING CENTRE

Published by Commonwealth Secretariat

ISBN 0-85092-550-9

Citation World Conservation Monitoring Centre. 1998. *WCMC Handbooks on Biodiversity Information Management. Volume 7: Data Management Fundamentals*. Reynolds, J.H. (Series Editor). Commonwealth Secretariat, London. ix + 33pp.

Typeset by Bookcraft Ltd, Stroud, Gloucestershire, England

Cover design Michael Edwards

Photography by J.S. Donaldson (Wood's cycad, *Encephalartos woodii*); D. & I. Gordon (Mali landscape; Plant study, Ghana; Thai forest; Rock hyrax, *Procavia capensis*); IUCN/J. McEachern (Diver and fish); WCMC (Ecoregion and Africa maps; GIS work; Workshop facilitation).

Available from IUCN Publications Services Unit
219 Huntingdon Road, Cambridge, CB3 0DL, UK
Tel: +44 1223 277894; Fax: +44 1223 277175
Email: iucn-psu@wcmc.org.uk
http://www.iucn.org/bookstore/

Contents

ACKNOWLEDGEMENTS

The generous support of the *United Kingdom's Darwin Initiative for the Survival of Species* has provided for the development of a comprehensive programme of training in biodiversity information management. This programme comprises an international training team, drawing on expertise from collaborating organisations around the world; the preparation of a training resource in the form of a handbook series and related materials; and the development of computer-based demonstration tools. Training is being promoted through the delivery of post-graduate modules, and through regional and national workshops which have received additional support from The British Council, British Airways Assisting Conservation Scheme, and contributions from participating organisations. The programme has been appropriately titled *Darwin Initiative Training in Biodiversity Information Management.*

Development of the handbooks has also benefited from experiences gained through the Biodiversity Data Management (BDM) Project, administered by the United Nations Environment Programme (UNEP) and funded by the Global Environment Facility (GEF), and related initiatives supported through the European Union (EU) and European Environment Agency (EEA). Indeed, Volume 6 draws extensively on one of the key outputs of the BDM Project, the *Guide to National Institutional Survey* (UNEP/WCMC 1998), developed in consultation with participating countries, the BDM Advisory Committee and the UNEP management team. The concept of an information cycle was developed in collaboration with the International Institute for Environment and Development (IIED) with support from the UK Department for International Development (DFID). The handbooks have been published through the generous support of the Commonwealth Secretariat.

Fundamental to the development of this programme have been the partnerships established with training organisations around the world. These organisations have worked collaboratively in hosting workshops, in reviewing the handbook materials, and in providing guidance on how regional and national training needs can be met most effectively. The training programme has significantly benefited from the input of numerous individuals working in the field of biodiversity information management. Among these individuals, particular mention goes to Professor Ian Crain and Gwynneth Martin of the Orbis Institute, Ottawa, Claire Appleby, an independent consultant, and to Drs Jake Reynolds and John Busby of WCMC for their insightful work in developing the handbook series. Thanks are also extended to Laura Battlebury for her tireless administrative and logistical support. The series

editor for the handbooks was Jake Reynolds, while Donald Gordon managed the overall project.

To the many individuals, both within and outside WCMC who have contributed to the development of materials and the delivery of training in biodiversity information management, a profound debt of gratitude is owed. It is through this collaborative effort that a service is being developed to contribute to the conservation and sustainable use of living resources.

BACKGROUND

The purpose of the *WCMC Handbooks on Biodiversity Information Management* is to support those making decisions on the conservation and sustainable use of living resources. The handbooks form part of a comprehensive programme of training materials designed to build information-management capacity, improve decision-making and assist countries in meeting their obligations under Agenda 21 and the Convention on Biological Diversity.

The intended audience includes information professionals, policy-makers, and senior managers in government, the private sector and wider society, all of whom have a stake in the use or management of living resources. Although written to address the specific need for improved management of biodiversity-related information at the national level, the underlying principles apply to environmental information in general, and to decision-making at all levels. The issues and concepts presented may also be applied in the context of specific sectors, such as forestry, agriculture and wildlife management.

The handbooks deal with a range of issues and processes relevant to the use of information in decision-making, including the strengthening of organisations and organisational linkages, data custodianship and management, and the development of infrastructure to support data and information exchange. Experience suggests that some of the greatest challenges in information management today are concerned with organisational issues, rather than technical concerns in the delivery of information which supports informed decision-making. Consequently, topics are addressed at management and strategic levels, rather than from a technical or methodological standpoint, and alternative approaches are suggested from which a selection or adaptation can be made which best suits local conditions. Nevertheless, in adopting this framework approach, we have tried to adhere to recognised conventions and formalisms used in information management and trust that in producing a 'readable' set of handbooks the integrity of the materials has not been compromised.

Overall, the handbook series comprises:

Companion Volume	
Volume 1	Information and Policy
Volume 2	Information Needs Analysis
Volume 3	Information Product Design
Volume 4	Information Networks
Volume 5	Data Custodianship and Access
Volume 6	Information Management Capacity
Volume 7	*Data Management Fundamentals*

Collectively, the handbook series promotes a shift from tactically based information systems, aimed at delivering products for individual project initiatives, to strategic systems which promote the building of capacity within organisations and networks. This approach not only encourages data to be managed more effectively within organisations, but also encourages data to be shared amongst organisations for the development of the integrated products and services needed to address complex and far-reaching environmental issues.

The handbook series can be used in a number of ways. Individual handbooks can be used to guide managers on specific aspects of information management; they can be used collectively as a reference source for strategic planning and project development; they can also provide the basis for a series of short courses and training seminars on key challenges in information management.

The companion volume provides the background to the handbook series. It also assists readers in deciding which handbooks are most relevant to their own priorities for strengthening capacity.

A second series of handbooks is planned to provide more detailed guidance on information management methodologies, including the areas of data and technology standards, database design and development, application of geographic information systems (GIS), catalogues and metadatabases, and the development of decision-support systems. The current series deals only briefly with formal system development methodologies, and for more detailed treatments the reader is encouraged to access the wide range of published and electronic resources available in libraries and on the Internet, some of which are alluded to in individual handbooks and reference sections.

A number of computer-based training tools have been developed to accompany the handbook series and are used in the training programme. These are based on a protected areas database, a tree conservation database, a GIS demonstration tool and a metadata directory. They aim to demonstrate key aspects in the collection, management and analysis of biodiversity data, and the subsequent production and delivery of information. They also illustrate practical issues such as data standards, data quality-assurance, data access, and documentation. Each training tool is supported by a user guide, together with a descriptive manual which traces the evolution of the tool from design, through development to use.

1 INTRODUCTION

Organisations which are assigned — and accept — responsibility for managing datasets are known as **custodians** (see Volumes 4 and 5). They will normally be regarded as being in the best, or most appropriate position to do so. The custodianship of **essential datasets** is especially important, since these are needed by many users for many purposes (see Volume 3). Examples include the basic demographic, geographic and biological data underpinning the analysis of human impacts on the environment.

The key requirement is to manage data in such a way that they can be readily converted into a variety of information products, for a variety of users, thus ensuring that they are flexible enough to respond to the demands of decision-making. This is a difficult challenge for custodians, but one which pays off with an **efficient information infrastructure**. The goal is to collect, store and quality-assure data just once, but access them many times for many different purposes (UK Government 1995).

In order to reap the benefits of efficient infrastructure, including lower costs and better services to users, custodians require certain basic **capacities**. These may need to be strengthened, perhaps in collaboration with other organisations, to ensure that the right balance of data, expertise, facilities, management systems and partnerships is available (see Volume 6). However, capacity alone does not breed efficiency; some fundamental insights and processes relating to the management of data must also be considered. These are summarised below and are examined in later sections.

- **Data flexibility**

  Data should be stored in their primary form, not classified, aggregated, or otherwise interpreted, so that they can be employed in the widest possible range of applications.

- **Data standards**

  Data should be collected, managed and distributed following agreed standards or conventions. This reduces transaction costs and facilitates comparison of results in space and time.

- **Data quality-assurance**

  Data quality — which is a measure of the **fitness for use** of a dataset for a specific purpose — may be assured through a number of processes, including data validation, documentation and protection.

- **Appropriate use of technology**

  Data should be managed within an environment that is conducive to data storage, processing and retrieval. Information and communication technologies are ideally suited to this task, and should be applied as appropriate and sustainable.

2 DATA FLEXIBILITY

Environmental data record phenomena in the physical environment. Some of these recordings are factual, for example the grid reference of the place where a species was observed, the dimensions of a tree, the weight of a log, the annual precipitation at a site, or the absorptive capacity of a soil profile. These are all **primary data** based on facts which can be measured against stable, widely accepted standards.

Secondary or **derived data** are obtained from primary data by a process of classification or interpretation, either at the time of measurement or later. Examples include species name, vegetation type, forest canopy extent and climatic zone. Derived data are not a substitute for primary data, and should not be stored permanently unless the primary data used to create them are also available. This is because derived data slowly degrade in value and, ultimately, become useless as concepts and paradigms shift. For example, if the only data on species distribution is an outline drawn on a map, this may become redundant if the species is split or otherwise disaggregated following a taxonomic revision. A better approach would be to store the primary data relating to the identification and location of the species in the field, so that new outlines can be derived as necessary.

Primary data are much more flexible than derived data. They can be used for a wider range of applications because they have not been modified for a specific function. For instance, daily rainfall measurements in millimetres from a local weather station can be used to assess local climatic fluctuations. They can also be fed into national or international climate monitoring programmes, or be integrated with other data to assess the capability of an area to support biodiversity. If the rainfall measurements had been classified at time of collection into, say, five secondary categories (e.g. very low, low, medium, high, very high), then this flexibility would have been lost, resulting in fewer potential applications for the data (such categories may be too coarse or meaningless in other contexts).

To ensure that data remain flexible, they should be collected and stored in their primary form, not classified, aggregated or otherwise interpreted. However, this rule does not need to be implemented rigidly; it may be subjected to intelligent assessment in each case. No one, for example, would refuse to store the names of species, even though they are susceptible to change. The process of deciding which type of data to store involves risk assessment. Given the high costs of collecting and managing data, the benefits of doing so needs to be balanced against the risk that data will become obsolete or prove to be inflexible. To assist with this judgement, Box 1 highlights the key characteristics of primary and derived data, and compares these with the information that both can be used to generate (see Volume 3).

It should be noted that perceptions of primary data and derived data vary considerably, according to the particular individuals or organisations concerned. For example, derived data to a scientific researcher may be regarded as primary data by a policy-maker, and be subjected to further analysis and interpretation. Despite differences in perception, the principle of storing primary data holds true within any particular domain, although between domains it may not.

Box 1 The nature of data and information

- **Primary data**

  These are facts which result from measurements or observations about the world, referenced to stable, widely accepted standards. The latter include absolute measures, such as units of length, volume or density.

- **Derived data**

  These are data obtained from primary data by a process of classification or interpretation, either at the time of measurement or later. They may be referenced to absolute measures but more commonly relate to professionally agreed conventions and products, for example maps which comply with an accepted structure and format.

- **Information**

  This is altogether different to data: it is the knowledge derived from the analysis, integration and interpretation of data, including 'expert' opinion. Unlike data, which may be applied to a range of purposes, information is produced for a specific purpose and has a short shelf life. Because of its transient nature, information should not be stored in databases unless this is judged to be cost-effective.

3 DATA STANDARDS

3.1 Overview

Standards enable people to communicate with each other in recognisable ways: languages are a good example. In the present context, data standards refer to agreed methods of collecting, managing and accessing data amongst a group of organisations. In the same way that language standards enable more efficient (and cheaper) communication, data standards enable more efficient use of data. The chief advantages of data standards are as follows:

- **Lower transaction costs**

  If data are available in standard formats, based on standard collection methodologies, users can absorb them more easily into their work. However, if standards are not applied then data may be perceived as incompatible, inappropriately focused or otherwise unusable. In summary, **lower transaction costs** are associated with accessing and using data when they are managed according to recognised standards.

- **Comparison of results**

  Without agreement on data standards, organisations tend to employ their own methods of collecting and managing data which, due to differences, complicate integration of the data with other sources at a later stage. Even within an organisation, methods may be applied inconsistently by different groups, or at different points in time. Data standards overcome this problem by enabling **comparison of results** in space and time, and between different sources.[1]

Admittedly, reaching agreement on data standards is a time-consuming, largely intellectual, activity requiring concrete and determined action to succeed. However, there is no other realistic way of reducing transaction costs or maximising the value

---

1 This is particularly relevant to the study of natural phenomena which, due to their incremental nature, tend to reveal themselves over long periods of time.

of expensively produced data. Data management is a resource-intensive activity, and it can be disappointing to discover that otherwise well-managed data are unusable due to lack of standardisation. This can be avoided by building in conformity from the early stages of a data management project, with the intention of widening the potential range of uses of data which are developed.

3.2 Types of Standard

Standards may be applied to all aspects of data management, from data collection and storage, to quality-assurance and distribution. They define accepted formats, structures, systems and procedures for managing data. Mostly, they define only minimum requirements, allowing those following the standards to exceed the requirements as appropriate. For example, a standard method of recording species' distributions might require observers to provide a location, date and species name for each observation. Observers would be welcome, however, to record any number of other factors in addition to this minimum set. Some of the potential range of data standards is described below.

- **Collection**

  Recording/measuring techniques for specific themes (e.g. biological records, human impacts, policy performance, sustainability); classification systems (e.g. soils, vegetation, climate, species names); criteria for assessing threats to biological resources.

- **Storage**

  Core data models/database structures; storage formats and media; methods of data retrieval; use of information technology; maintenance procedures.

- **Quality-assurance**

  Validation, maintenance and security procedures; documentation formats.

- **Distribution**

  Product definitions (e.g. map keys, acknowledgements, symbols); reporting formats; data transfer formats (interchange standards); protocols for electronic communication of files.

It is not always feasible or even desirable for organisations to adopt every type of standard. They may have their own, highly effective ways of managing data which could become compromised, possibly disrupted, by the blanket introduction of new and unfamiliar standards. Where increased efficiency is unlikely to follow the introduction of standards, they should not necessarily be pursued.

There is one group of standards which will almost always bring efficiencies, at very little cost. These are **interchange standards**, whose purpose is to streamline the transfer of data. The introduction of interchange standards has very little impact on the way organisations collect and manage their data internally, but has a strong impact on data mobility. Interchange standards focus mainly on the formats and media in which data are transferred. Because they apply solely to the distribution of data, interchange standards are simpler and cheaper to implement than wider-ranging standards.

A number of interchange standards already exist for the transfer of biological and geographic data. For example, the International Working Group on Taxonomic Databases for Plant Sciences has developed an International Transfer Format for Botanic Garden Records (Hollis and Brummit 1992). Interchange standards also exist to regulate the transfer of spatial datasets, which are often highly complex due to the varied nature of the data.[2] Most of these are based on the formats developed by the manufacturers of geographic information systems (GIS), for example the export formats of ARC/INFO and AutoCAD software. Such standards are privately controlled (i.e. they are **proprietary**) and may not necessarily reflect the needs of data users.

Non-proprietary standards have been developed to address this concern, although they are not yet in widespread use. For example, the Spatial Data Transfer Standard

---

2 For example, raster data, vector data, three-dimensional data and attribute data.

(SDTS), which is coordinated and promoted by the United States Geological Survey (USGS), consists of specifications for the organisation and structure of digital data transfer, definitions of spatial features and attributes, and encoding instructions for data transfer (Wortman 1992). The SDTS was approved by the United States Department of Commerce in 1992 (NIST 1992), and has been adopted by several other countries.

3.3 Development of Standards

Information networks provide an opportunity to reconcile existing standards — and agree new ones — in the interests of mobilising data for collective goals. Their development can be facilitated by one or more technical teams, arranged by the network's hub, who are tasked with reviewing and agreeing standards covering essential themes, and for publishing them for use by the network's partners (see Volume 4). Data standards are so important to a network that they cannot be overlooked, taken for granted, nor left to specialists who do not fully represent the network's interests.

Recognising that progress towards formally accepted data standards can be slow, organisations often develop their own, interim standards. The latter, sometimes referred to as *de facto* standards, are commonplace across many scientific themes, often having arisen to suit particular data collection and management priorities. Wherever possible, interim standards should build on existing standards within their theme, rather than risk duplication. For example, international initiatives have so far proposed at least seventeen definitions of sustainable forest management, many of which could be translated into national standards for forest monitoring (WBCSD 1996).

As the profile of interim standards is raised, and increasing numbers of organisations begin to adopt them, they may be vetted by the organisations concerned and formalised through publication. A good example of this process is the East African Biodiversity Network, now in its seventh year, which has successfully developed biological recording standards at a regional level. The network, which brings together biodiversity professionals from Kenya, Tanzania and Uganda, was originally established "in response to the need for biologists and conservationists to develop compatible working systems, from the database formats themselves to having common lists of scientific names" (NMK 1995).

A number of working groups were set up to develop data standards appropriate to the region. Taxonomic working groups are developing checklists and other standards for birds, mammals, reptiles, amphibians, fish, aquatic invertebrates and plants. A working group on important bird areas is responsible for listing, prioritising and surveying key sites; other groups are developing a regional gazetteer, habitat classifications, database structures, and policies on data exchange and training.

At the international level, the International Working Group on Taxonomic Databases for Plant Sciences (TDWG) was established by the International Union of Biological Sciences in 1985 to explore standardisation and collaboration between major taxonomic databases (Hollis and Brummit 1992). The group brings together all the major working taxonomic databases into a loose confederation. Through a series of international workshops, TDWG has developed a number of standards including the International Transfer Format for Botanic Garden Records (ITF) and a World Geographical Scheme for Recording Plant Distributions. The latter provides four nested levels comprising continents, regions, countries and botanical recording units.

4 DATA QUALITY-ASSURANCE

4.1 Overview

Data quality is a relative term, for which there are no absolute measures. In practice, data quality is a measure of the **fitness for use** of a dataset for a specific purpose, and cannot be determined before that purpose is known. For example, a topographic map at a scale of 1:500,000 might be considered 'high quality' for national-level planning purposes, but 'poor quality' for local planning. Thus, the quality of a dataset is clearly affected by its accuracy and validity, but is not necessarily defined by it.

The complexity of natural phenomena means that many environmental measurements are uncertain or subject to error. For example, it is inevitable that some species will be mis-identified in a large-scale biological inventory, even if the highest professional standards are employed. Similarly, the inference of vegetation categories from remotely-sensed satellite imagery will never be 100 percent accurate. Such uncertainties may or may not be a cause for concern, depending on the intended use of the data. Box 2 distinguishes three common forms of deficiency in environmental datasets which may affect data quality.

Recognising that most environmental datasets contain deficiencies, it is vital for custodians to pass on an understanding of these when a dataset is distributed for external use — otherwise users may not be able to derive the maximum benefit from it. Clearly, a description of known deficiencies is only one item of information required by users to employ the dataset fully and safely. Others issues to document relate to the accuracy of the data, the standards which have been followed, and the processing techniques which have been applied (see Section 4.5).

Procedures aiming to improve the quality of a dataset can be applied from the moment it is collected through to the time that it is distributed for use. These procedures, which are collectively known as **quality-assurance procedures**, are designed to satisfy the needs and expectations of users.

4.2 Quality-assurance Procedures

Quality-assurance refers to the overall process governing the quality of a product, from the time that it is originated to the time that it is used. In the present context, the

Box 2 Forms of deficiency in environmental datasets

- **Limitations**

  Limitations are structural deficiencies in a dataset which become clear when it is used for purposes other than originally intended. A good example is the use of a map with an inappropriate scale.

- **Uncertainties**

  Uncertainties are introduced when variables are measured against a non-objective standard, for instance when an area is classified as belonging to a particular habitat type which, itself, may be poorly defined.

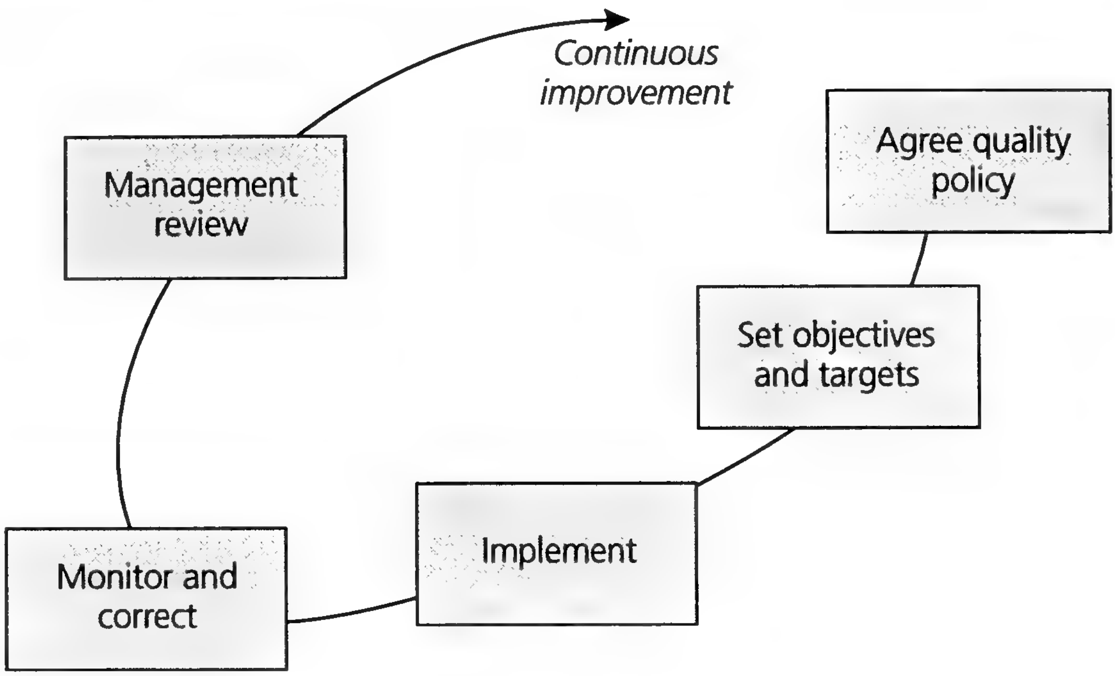
- **Errors**

  Errors are introduced when variables are measured incorrectly against an objective standard, for instance when the depth of a lake is recorded with the digits in the number accidentally transposed, or with the wrong units (e.g. feet instead of metres).

process begins with data collection and ends with distribution of information to users. Quality-assurance procedures can be applied during all stages of this cycle. These include procedures to validate, maintain, document and secure data. It is the responsibility of custodians to ensure that these procedures are implemented in line with accepted standards and user demands (see Volume 5). Policies, judgements and decisions all depend on them doing so.

Within an organisation, quality-assurance procedures should be defined within a **quality policy** that is well understood by appropriate staff. The policy should set challenging objectives and targets for staff to achieve, such as specific levels of numerical or spatial accuracy in data collection, allowable error rates during validation, or consistent standards of documentation. The targets need to be consistently applied across the organisation and be measurable for monitoring and review. As well as internal review, organisations should also seek feedback from users of its products and services. The combination of internal and external reviews allows the organisation to correct deficiencies in data quality and continuously improve its quality-assurance procedures. Figure 1 illustrates the essential steps of the quality-improvement process (adapted from BSI 1994).

Figure 1 Quality-improvement loop



4.3 Validation

Uncertainties and errors are introduced into a dataset in the natural course of data collection. The aim of validation is to eliminate these completely or reduce them to a background level where they do not interfere with the use of the data. Validation can be a labour-intensive and tedious task, but it is nevertheless a critical quality-assurance procedure. Key activities include:

- testing the accuracy and reliability of data prior to storage; and
- introduction of tools and methods to regulate data entry.

Basic tests should be run on data items before they are permanently stored (e.g. before new data items are added to existing datasets). These enable suspect or unusual data items to be identified and brought to the attention of experts for independent assessment. Box 3 describes some basic tests applied to species distribution records prior to inclusion in a large national dataset in Australia (Chapman and Busby 1995). Another good example of the expert assessment process is the validation of bird distribution records in East Africa. Here, national experts validate the vast majority of bird distribution records generated by field survey activities, but very unusual records are processed at the regional level by the Ornithological Sub-committee of the East Africa Natural History Society (Reynolds *et al.* In press).

Box 3 Example validation procedures for species dataset

- Records checked to see that all required data fields are present.
- Scientific names checked for validity.
- Grid references of terrestrial species checked for being over land, not water.
- Presence of a species in a certain location tested against a prediction based on bioclimatic factors, and outliers selected for further investigation.

Errors can be introduced into a dataset when it is stored, for instance in a computer. Common errors include the entry of incorrect numbers into a spreadsheet or incorrect boundaries into a map. As an illustration, take the entry of species data into a computer database. Suppose that a particular data entry screen has 10 fields (e.g. family, genus, species, common name, threat category, etc.), each taking, on average, 8 characters to fill. If the success rate of the typist is 99 percent, then the probability of the whole screen being completed correctly is, surprisingly, only 45 percent.[3]

3 If the probability of a single character being typed correctly is 99 percent (0.99), then the probability of 10 fields, each with 8 characters, being typed correctly is $0.99^{(10x8)} = 0.45$, which is 45 percent.

Such errors result largely from lack of care and attention by human operators, and training will help to reduce these. However, they can be reduced even more effectively through the introduction of tools and methods to regulate data entry. These promote consistency and enable operators to identify errors at the earliest detectable moment, so that they do not propagate or become buried in large volumes of other data.

A key feature is automatic validation, which involves performing 'reasonableness' checks on data items as they are entered, such as the geographic feasibility of a grid reference or the physical possibility of a particular measurement. Unreasonable values (e.g. a land-based animal observed at sea) can then be reported to the data entry operator, who can correct simple mistakes or seek expert advice as required. Even more effective at reducing errors are tools which allow the operator to select values from a set of pre-defined choices, eliminating the possibility of typographic errors completely. Automatic validation is especially useful in situations where consistency of data entry cannot be guaranteed, for instance when data are entered into large datasets by many different staff.

4.4 Maintenance

Most datasets become obsolete if they are left unmanaged for long periods of time. Measuring techniques may be improved, leading to more accurate and reliable data collection; new standards may be agreed, meaning that old structures and assumptions are no longer acceptable; and new formats, media and technologies may be evolved to manage data more efficiently. Unless a dataset is actively maintained, it may simply be overtaken by events leading to a gradual reduction in its usefulness. Key activities include:

- keeping it up to date;
- making sure it is kept abreast of significant standards; and
- adapting its structure, format and storage medium in line with user's needs.

Keeping data up to date involves establishing a routine for continuous, or at least regular, **enrichment of a dataset with new data**. Many projects fail to take account of this, with datasets being created to serve only immediate project objectives, rather than long-term capacity needs. This is inefficient, since new projects may have to build similar datasets from scratch. One of the distinguishing characteristics of a professionally-managed dataset is that it is maintained not only for immediate uses, but also for other applications — now or in the future — which could potentially benefit. As with other strategic approaches, this can create a funding challenge in the short term.

Earlier sections revealed the importance of data standards. These also evolve over time as new opportunities for standardisation are created through information networks and individual partnerships between organisations. Where relevant standards exist, they may be applied to datasets in order to ensure consistency and reduce transaction costs; where they evolve, datasets should evolve with them to maintain these advantages.

Over time, increasing numbers of users may apply a dataset to their tasks. Feedback from users, for instance their impressions of the strengths, weaknesses and overall usefulness of the dataset, can be used to adapt the structure, format and medium in which it is made accessible. Note that the dataset itself can be **managed** in whatever form is discovered to be most efficient by the custodian, but it should be made **accessible** in the form which is most acceptable to users (see Section 3.2).

The opportunities created by rapidly-changing information technologies, storage media and low-cost communications, impose a continuous challenge on those attempting to maintain datasets. However, it is far more important for data managers to maintain the content of their data than worry about keeping up with the latest technology; from a user's perspective, all that is required is a simple and cheap source of quality-assured data.

4.5 Documentation

When a dataset is released to an external user, knowledge of its limitations, uncertainties and errors is lost unless this understanding is passed on in the form of documentation. As well as knowledge of its deficiencies, users may require a host of other items of information in order to employ the dataset fully and safely.

In the past, custodians rarely devoted much attention to documenting their datasets. This was because the latter were usually built for one specific project by people who well understood the nature of the data, including its deficiencies. At the end of the project the data were archived, filed or neglected. Today, however, datasets may be used many times for many purposes, and documentation is regarded as a strategic asset enabling custodians to maximise the value they derive from a specific data source. One of the driving forces of this change is the growth of information networks which depend on organisations being granted simple and cost-effective access to data.

In summary, custodians document their datasets for two important reasons:

- to increase internal effectiveness by clarifying the function and quality of their datasets; and

- to facilitate use of their data by others.

Box 4 lists some potential aspects of a dataset to document. The fundamental principle to follow is **truth in labelling**. This means that the dataset should be exactly as described and of a quality which is suitable for its stated and implied uses. Assessments of the completeness and accuracy of documentation should be undertaken periodically, especially in the case of essential datasets (see Volume 3), preferably by an independent auditing team.

4.6 Data Security

A range of operational procedures are necessary to guarantee the security of a dataset. This applies whether or not data have been computerised. Indeed, if they are not in electronic form, then it may be considerably more difficult to manage them securely.

In general, threats to electronic data security tend to be greatest where the physical environment is hostile to computing equipment (e.g. extremes of temperature, high humidity or dust), where electronic interference is strong (e.g. in hospitals, industrial plants, locations near transmitters), where power supplies are uneven or unpredictable, and where informal and therefore virus-prone computer networks are the primary means of data transfer.

The most important requirement is to protect data from accidental erasure, which may occur due to human error in copying and reorganising files, updating records or other 'maintenance' procedures. Erasure may also occur due to mechanical failure of disk drives, or logical faults caused by power failures or fluctuations. Computer viruses also pose a threat to data security, although this is often greatly over-estimated (they certainly are a nuisance however).

Box 4 Aspects of a dataset to document

- Title/theme.
- Contact details of custodian.
- Intended/unwise/improper uses.
- Accuracy/resolution/scale.
- Data collection methodology (or original sources of data).
- Data structure/model.
- Data management standards followed.
- Processing and interpretation techniques applied.
- Known limitations, uncertainties and errors.
- Currency of data.
- Life expectancy (e.g. date of next update).
- Quality-assurance procedures applied.
- Quality targets.
- Access conditions/procedures/costs.
- Available formats and media.

Box 5 describes a number of protective measures which help to combat threats to data security. Such procedures can be elaborated within the overall quality policy of the organisation, or be prepared separately in the form of an operating manual. Specific plans to cope with emergencies should also be considered, for instance hardware malfunction, fire or theft. Organisations should accord a high profile to data security. On occasion, an entire project or programme has been forced to close due to loss of essential data. This occurred once in the South Pacific when a freak wave struck the office of a custodian, eliminating its data. No copy of the data was maintained off-site.

Box 5 Procedures for protecting data

- Regular (daily, weekly and monthly) backup of all critical data on removable electronic media (magnetic tape or optical disk).
- Storage of backup media off-site (away from the workplace) in order to restore data after damage or theft of key equipment.
- Periodic test restoration of backed-up data to ensure that the procedure is effective.
- Periodic test recovery from simulated virus attack, hardware malfunction or other disaster.
- Regular virus-checking with up to date software.
- Avoidance of unlicensed or borrowed software, computer games or other personal software.
- Power regulation via the use of uninterruptable power supplies, surge protectors and radio interference filters.

5 USE OF INFORMATION TECHNOLOGY

5.1 Overview

If applied in an appropriate and sustainable manner, information technology can lead to considerable cost savings and efficiencies in an organisation. Alternatively, if technology is allowed to dictate strategy, costs are likely to rise and existing work patterns may be disrupted. Such situations demand a fundamental re-appraisal of the role of information technology. In essence, information technology should **support, not drive** data management objectives.

Although data can be managed without modern information technology, the latter has some important advantages over manual techniques. For instance, computers can be used to store large volumes of data and perform very rapid and complex analyses. They can also be used to validate data as they are entered and be used to produce multiple and varied reports from the same data. These advantages widen the range of purposes to which the data may be applied.

Information technology also brings certain disadvantages, particularly in the form of additional **complexity and costs**. Almost every item of new technology brings with it a maintenance, support and training overhead. Box 6 highlights several situations which, on balance, would benefit from the appropriate use of information technology.

5.2 Selecting Technology

In any given situation, the best type of information technology is that which is most appropriate to the tasks at hand — both now and in the future. In particular, the issues of scalability, connectivity, compatibility and sustainability need to be closely examined (see Box 7). Following this analysis, the advantages and disadvantages of different technological solutions should be tested under realistic local conditions before procurement takes place. Although useful, manufacturers descriptions, magazine reviews and specialist information services (e.g. Internet newsgroups and bulletin boards) should not be relied upon for strategic procurement decisions.

It should be noted that some characteristics of information technology are subjective, such as the ease of use of a software package or the quality of a scanned

Box 6 Situations where information technology could be beneficial

- Data contain relationships which are too complex, or are too great in volume for the capabilities of manual filing systems or word processors.
- It is necessary to integrate data from several sources into a combined output.
- There is a need for the data to be shared amongst more than one user in a single organisation, or with other organisations.
- Data require extensive searching, sorting or updating.
- Frequent and varied reporting of the data is required.

image. Thus, selecting technology purely from a list of features is unlikely to be satisfactory. Like before, a real-life test is the best way of determining whether technology will be suitable under the expected working conditions.

A wide range of options exist for managing data. These include single (stand-alone) computers running local copies of data-management software; locally-networked computers with shared software running on a file server (i.e. a Local Area Network or LAN); client-server architectures[4] which integrate the best characteristics of personal computers (friendly software and quick response) with the best traits of file servers (high storage capacity, fast data processing, good security); and fully-distributed databases consisting of a series of remote computers linked via permanent or dial-up communication lines (i.e. a Wide Area Network or WAN). The decision as to which option to select should be taken after a thorough examination of the factors summarised in Box 7. Clearly, the nature and extent of the data to be stored will influence this decision greatly, as will the degree to which the data need to be accessed electronically by internal and external users.

---

4 This option is becoming increasingly popular for medium- to large-sized organisations relying heavily on data management for their core business.

Box 7 Considerations when selecting information technology

- **Scalability**

As the number of users, records or attributes grow, an application that once performed well on a low-cost computing architecture can deteriorate in performance quickly. Typically, stand-alone or small network computer architectures are most likely to suffer from this problem, which explains the rise of more sophisticated architectures, such as client-server.

- **Connectivity**

To enable rapid exchange of data between individuals and organisations, electronic connectivity is desirable. This could take the form of a group of locally-networked computers sharing a common storage area, or more sophisticated dial-up communication lines to external services, such as the Internet and private networks. The capacity to connect computers together into more powerful resources is increasingly recognised as the key to rapid access and use of data.

- **Compatibility**

The issue of compatibility is diminishing as manufacturers evolve a range of standard specifications for their IT products. However, the specifications — which are often proprietary in nature — are still too varied and numerous to discount the problem entirely. As far as computing platforms are concerned (i.e. computer hardware plus operating system), major decisions include whether to adopt IBM-PC compatible computers which running derivatives of the Microsoft Windows operating system, or larger workstations running the UNIX. Since the technologies are changing so rapidly, there is really no 'best' solution other than to adopt a platform which has proved to be reliable and useful in circumstances similar to those anticipated, working on the principle that, in such cases, compatibility issues are unlikely to cause serious disruption.

continued overleaf

Box 7 Considerations when selecting information technology (cont.)

- **Sustainability**

  For information technology to deliver long-term improvements in effectiveness, sufficient funds and expertise must be available for users to exploit its potential fully and not be disadvantaged by its costs in terms of training, technical support and maintenance. Technology which has proven effective under the prevailing conditions is usually the wisest choice.

One of the most common forms of software used to manage environmental data is the relational database management system (RDBMS). These offer flexibility and performance at modest cost, although they are not designed to manage large-scale textual sources (these are more effectively managed in a word-processing package). Other key software include geographic information systems (GIS), which store, integrate and analyse spatially-referenced data, and tools such as spreadsheets, statistical packages and special-purpose environmental modelling software.

5.3 Database Development

Database development involves designing and building the systems necessary to manage one or more related datasets. Generic methods have been proposed to develop databases, and the ideas presented in following paragraphs attempt to simplify and summarise these. For clarity, database development is partitioned into two phases: database design and applications development.

- **Database design**

  This involves identifying the structure and functionality of the database. The required sources of data are made clear, and the integration and processing techniques needed to achieve the desired outputs are identified. The design process gives rise to a **functional specification**, which is independent of both hardware and software, and does not assume any particular method of physical data organisation (in practice, the technology available — which may be constrained by budgetary limitations — may affect the design of the database).

An important part of the design process is **data modelling**. This is the analysis of data objects and the identification of the relationships among these data objects. A common approach is to use **entity-relationship (E-R) diagrams**, as developed by Chen (1976). Quite simply, an E-R diagram depicts the contents of a database: an **entity** (shown as a rectangle) is an object (or 'thing') about which data are collected; and a **relationship** (shown as a line) shows the connections between the entities. The nature of the relationships between the entities indicates the number of occurrences of one entity that may be associated with a single occurrence of the other.

Three types of relationship are possible: one-to-one, one-to-many and many-to-many. For example, the entity 'Protected Area' may contain data such as the protected area names, legal status, size and so on. There may be a relationship to a 'Country' entity, indicating that each protected area is located within one or more countries (i.e. a one-to-many relationship). Although protected areas normally fall within the borders of a single country, being aware that it is possible for them to straddle more than one country has important implications for design. Indeed, failing to establish the correct relationship at an early design phase could restrict the development of the application at a later date. As discussed in Section 2 of Volume 2, it is vital to identify problems in the design phase, before large investments have been made in implementation.

In summary, the design process provides:

- ✓ a stable base from which to coordinate the development of the database, including the selection of appropriate equipment for implementation;
- ✓ a conceptual model which is free of implementation considerations, and which can be used as a point of reference when adding to or modifying the functionality of the database; and
- ✓ a baseline from which an optimum physical data organisation can be produced.

- **Applications development**

This involves creating a fully-functioning database using the data management software selected for implementation (see Section 5.2). Entities in the database design become **tables** in the software, and attributes become table **fields**. The way in which relationships between the entities are dealt with depends on which software is used; if it does not support some types of data relationship, then this has to be resolved by altering the database design.

Each field in the database is documented in terms of its purpose, data type, size and order in its corresponding table. When pooled across all the tables of the database, these definitions are known as the **data dictionary** of the database, and provide a description of its content, format and structure.

After the database tables have been created, they are **populated** with data. If the data are already computerised, this may be achieved by directly importing them into the database, plus associated re-structuring and formatting. If the data are only available in hard copy form, they will need to be entered manually into the database via the keyboard or, in the case of maps, images and structured text, via other input devices, such as scanners, digitising tablets and related software.

Most data management software packages enable developers to customise data entry procedures, for example by enforcing certain formats and validating or correcting data items as they are entered. This concept can be extended to other procedures, such as the querying and reporting of data, and saving data to removable media (e.g. a floppy disk) for back-up or delivery to users. The combination of data entry, querying and reporting features, security features and, of course, the underlying data tables, is known as a database **application**.

Database applications need not be created perfectly at the first attempt. Indeed, there is an advantage in developing **prototype** applications over a short time frame, and at low cost, in order to provide users with a means of refining their needs from the database. Prototyping was discussed in Volume 3 in the context of information product development, but it is equally useful and, perhaps, more essential during the development of databases. The aim is to allow problems to be identified and corrected early on in the database development process, circumventing costly modifications at a later stage.

6 CASE STUDY: TREE CONSERVATION DATABASE

6.1 Overview

With support from the Government of the Netherlands, the World Conservation Monitoring Centre (WCMC) and the Species Survival Commission (SSC) are working closely with a range of other national and international organisations to develop a global information service on the conservation and sustainable-use of trees. Reliable and up-to-date information on the distribution, conservation status, local uses and economic values of trees is a priority requirement for the planning of sustainable forest management and biodiversity conservation.

The Tree Conservation Information Service aims to be of value to individuals and organisations whose decisions rely on access to high quality data and accurate information. Whether determining the best use of local land or negotiating the obligations of an international treaty, authoritative data and information on tree species will inform the process and increase the likelihood that sustainable practices are employed and negative environmental consequences are minimised.

The service is underpinned by a **Tree Conservation Database**, developed with the following objectives:

- to enable the collation of data on the distribution, conservation status, local uses and economic values of tree species worldwide;
- to provide a low-cost software tool for the management and reporting of these data; and
- to provide the basis of a tree conservation information service on the Internet.

The database will be distributed to users in electronic form to enable storage, editing, analysis and reporting of tree-related data. It will also be analysed centrally to produce outputs such as a **World List of Threatened Trees** (using the new IUCN threat categories).

6.2 Target Audience

The Tree Conservation Information Service aims to help governments make informed and justifiable decisions on the conservation and sustainable use of trees. In addition to national governments, the information service will serve the needs of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), the International Tropical Timber Organisation (ITTO), the Convention on Biological Diversity (CBD), the Forest Stewardship Council (FSC), industry and other groups.

6.3 Information Needs Analysis

Before building the information service, considerable time and effort was invested in consulting with prospective users and collaborators (see Volume 2). Early in the project, a tree and timber database questionnaire was prepared and posted (July 1995) to over 500 organisations, representing national governmental forestry and conservation departments, bilateral and multilateral development agencies, national and international non-governmental organisations, research organisations, forest product trade organisations, and individual experts. The questionnaire had two main aims:

1. to identify priority needs for the Tree Conservation Information Service; and

2. to determine the availability and quality of existing data sources.

Following this exercise, the information needs analysis became more interactive. Scientific and technical experts were brought together at a workshop, which provided further opportunity to identify key data requirements and the types of information products the service could provide. A range of possible questions which the service could shed light on, and which could be addressed potentially by the Tree Conservation Database, are listed in Box 8.

Following this consultative process, the broad categories of data to be included in the database were analysed (see Volume 3). These included data on taxonomy, species distribution, conservation status, local uses, trade, threats, legal protection, contacts and other data sources.

Box 8 Potential issues to be addressed by the database

- Is the species of conservation concern?
- Has the species been evaluated for the new IUCN threat categories?
  - if so, what is the category and criteria by which it was assigned?
  - what information is available to support the threat category?
- What is the distribution of the species?
- What are the uses of the species?
- Is the use of the species sustainable?
- What are the current levels of trade in the species?
- What are the types, levels and values of use that are being made of the species?
- Is the species legally protected – regionally, nationally, internationally?
- What are the administrative and legislative structures pertaining to the conservation/sustainable use/management of tree species in any particular context?
- What are the implications of specified human actions and/or natural phenomena?
- What current actions are being taken to manage tree species, and how effective are they in achieving their objectives?
- Which individual or organisation holds, has access to, or can generate the data or information relevant to a specific issue?

6.4 Database Design

On the basis of the information needs analysis, and on-going discussions with other organisations and SSC specialist groups, a functional specification was developed for the database. This included an entity-relationship (E-R) diagram, table and field descriptions, plus a description of all the required functions and outputs of the database. The E-R diagram, which illustrates the links between the main data tables and look-up tables, is illustrated in Figure 2.

In terms of functionality, the main outputs (products) of the Tree Conservation Database were conceived at an early stage. These range from simple list-type reports, to more complex fact-sheet summaries and statistics. To provide flexibility within the database, there are comprehensive user- defined reporting capabilities. Standard reports and statistics include:

- species lists by distribution and/or threat category;
- species summaries by taxonomic group and/or distribution and/or threat category;
- total number of species in each threat category by taxonomic group and/or distribution; and
- total number of endemic species by taxonomic group and/or distribution.

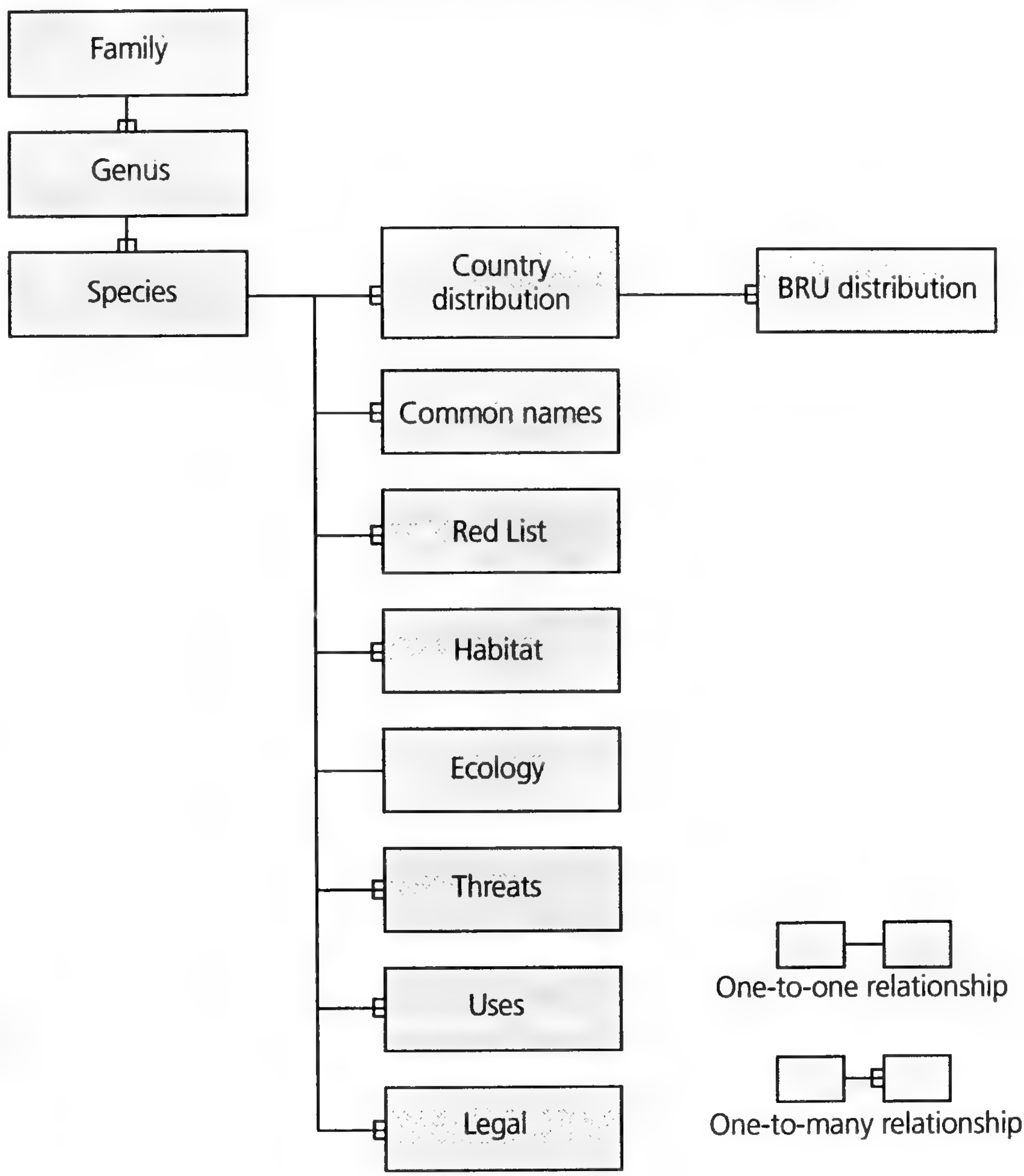
Each report can be printed and/or saved as a text file (and then, for example, used in a word processor) or as a delimited file (and then, for example, used with a spreadsheet). Reports may be written to file in many standard formats, including plain text, delimited text, dBASE and Excel.

6.5 Applications Development

A prototype database application was developed using data management software (RDBMS) familiar to WCMC staff. The prototype enabled an interactive approach to database development, and ensured that the final database correctly met user expectations. It was built quickly and cheaply, and served to focus attention on user requirements at project workshops.

Following review of the prototype, it was possible to make final decisions on the database design, plus the hardware and software to be used for implementation. The

Figure 2 Entity-relationship (E-R) diagram for the database



selection of hardware and software was also guided by the need to link effectively with other applications, including geographic information systems (GIS). In addition, it was necessary to run the database in single-user and multi-user environments, and to ensure compatibility with the latest generation of Windows-based network environments, especially Windows 95 and Windows NT.

6.6 Data Standards

The database employs standards in three main areas, as described below:

- **Taxonomy**

  Through the use of look-up tables, only valid entries of family and genus name are permitted, according to Brummitt (1992). In addition, the inclusion of scientific authority aids identification of particular species.

- **Geographic areas**

  Country names follow those specified by ISO (International Organization for Standardization) in standard ISO 3166 (codes for the representation of names and countries). At the sub-national level, areas are named in accordance with the internationally-agreed Basic Recording Units (BRU), described by Hollis and Brummitt (1992) and endorsed by the Taxonomic Database Working Group (TDWG).

- **Threat categories**

  The IUCN categories and criteria (adopted by the IUCN Council on November 30, 1994) provides a system for classifying the conservation status of species on a global scale. Species are evaluated and classified into one of eight categories: Extinct (EX), Extinct in the Wild (EW), Critically Endangered (CR), Endangered (EN), Vulnerable (VU), Lower Risk (LR), Data Deficient (DD), Not Evaluated (NE). The criteria by which the categories are applied are specified for each species.

6.7 Data Quality

The project's approach to data quality-assurance is essentially an *ad hoc* process, complemented by more thorough, structured reviews. In particular, the Tree Conservation Database allows data to be added and updated continuously, to reflect on-going changes in the assessment of the conservation status of species. As environmental conditions change, and as new research broadens the knowledge-base, then such re-evaluations are often required. The Tree Conservation Database caters for new assessments and new data by allowing modifications to be made, whilst retaining the original data.

Effective back-up is a further important issue addressed by the database. Although users and organisations may have their own back-up procedures, it was felt necessary to provide a special-purpose back-up option within the database, to compliment these other processes.

In some cases, access to the database may need to be controlled and, for this reason, password entry is included. Once a user has successfully logged into the database, the functions available to them are also determined by their privilege settings, of which three are defined as follows:

- basic user (view only; no access to database administration tools);
- user (add and edit data; no access to database administration tools); and
- administrator (all functions).

6.8 Cooperation and Partnership

Cooperation with other organisations is an important feature of the *Tree Conservation Information Service*, aimed at maximising the contribution of the project to related initiatives. For example, important partnerships were developed between WCMC and IPGRI and between WCMC and FAO, relating to the following initiatives, respectively:

- *REFORGEN* database system. This global database (developed by the Forest Resources Division of FAO) is designed to house information related to the world's forest genetic resources.
- *TREESOURCE*. This global information system on forest genetic resources represents a collaborative effort between FAO, the Centre for International Forestry Research (CIFOR), the International Center for Research in Agroforestry (ICRAF) and IPGRI, and has been designed to provide readily reliable and accessible information on forest genetic resources.

7 REFERENCES

Brummitt, R.K. 1992. *Vascular plant families and genera*. Royal Botanic Gardens, Kew, London, UK.

BSI 1994. *Draft BS ISO 14001: Environmental Management Systems – Specification with guidance for use*. Technical Committee ESS/1 – Environmental Management Systems. British Standards Institute, London, UK.

Chapman, A.D. and Busby, J.R. 1995. *Linking plant species information to continental biodiversity inventory, climate and environmental monitoring*. In: Mapping the Diversity of Nature. Miller, R.I. (Editor). Chapman & Hall, London, UK.

Chen, P.P. 1976. The Entity-Relationship Model – towards a unified view of data. *ACM Transactions on Database Systems*, Vol. 1 No. 1, pp 9-36.

Hollis, S. and Brummit, R.K. 1992. *World geographical scheme for recording plant distributions*. Plant Taxonomy Database Standards No. 2 v.1.0. Hunt Institute for Botanical Documentation, Carnegie Mellon University, Pittsburgh, USA.

NIST 1992. *Federal information processing standard publication 173*. National Institute of Standards Technology, United States Department of Commerce.

NMK 1995. *Biodiversity in East Africa: Proceedings of the sixth regional workshop on biodiversity databases*. National Museums of Kenya, Nairobi, Kenya.

Reynolds, J.H., Tushabe, H. and Kasoma, P. In press. The Establishment of a National Biodiversity Data Bank for Uganda. *Journal of East African Natural History Society*. East African Natural History Society, Nairobi, Kenya.

UK Government 1995. *Biodiversity: The UK Steering Group Report. Volume 1: Meeting the Rio Challenge*. The Stationery Office Ltd (HMSO), London, UK.

WBCSD 1996. *A Changing Future for Paper: An independent study on the sustainability of the pulp & paper industry*. World Business Council for Sustainable Development, Geneva, Switzerland.

Wortman, K. 1992. *Spatial Data Transfer Standard: The Key to Data Sharing.* U.S. Geological Circular, 1992 <http://www.usgs.gov/reports/yearbooks/1992/nmp_spatial.html>

WCMC Handbooks on Biodiversity Information Management

These handbooks have been developed for use by senior decision-makers and mid-career professionals. They review the issues and processes involved in the management of biodiversity information to support the conservation and sustainable use of living resources. They also provide a framework for the development of national plans and strategies and for meeting reporting obligations of international programmes and conventions. Collectively, the handbook series may be used as a training resource or, more generally, to support institutions and networks involved in building capacity in information management.

Companion Volume
Volume 1 Information and Policy
Volume 2 Information Needs Analysis
Volume 3 Information Product Design
Volume 4 Information Networks
Volume 5 Data Custodianship and Access
Volume 6 Information Management Capacity
Volume 7 Data Management Fundamentals

DARWIN INITIATIVE SECRETARIAT
4-A2 Ashdown House, 123 Victoria Street
London SW1E 6DE. United Kingdom

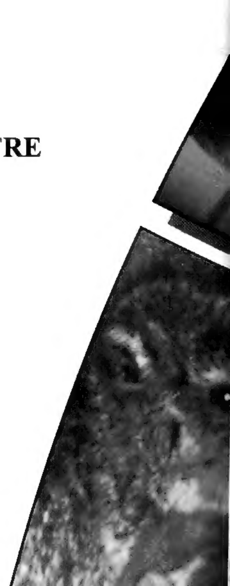
COMMONWEALTH SECRETARIAT
Marlborough House, Pall Mall
London SW1Y 5HX. United Kingdom

WORLD CONSERVATION MONITORING CENTRE
219 Huntingdon Road
Cambridge CB3 0DL
United Kingdom

Tel: +44 1223 277314
Fax: +44 1223 277136

E-mail: info@wcmc.org.uk

WCMC Internet Home Page
http://www.wcmc.org.uk



ISBN: 0-85092-550-9